Normalized Unscented Kalman Filter and Normalized Unscented RTS Smoother for Nonlinear State-Space Model Identification

Masaya Murata† and Kunio Kashino†‡

Abstract—A Kalman filter (KF) and Rauch-Tung-Striebel smoother (RTSS) provide the minimum mean square estimates of states for linear state-space models with additive Gaussian system and observation noises given a series of the past, current, and future observations. When the noise statistics such as the variances are unknown, initially normalizing the KF and RTSS algorithms by the total of the unknown variances provides new state estimation algorithms. We call these algorithms the normalized KF and normalized RTSS and on the basis of the log-likelihoods scored by the multiple trials with varied parameters, we can effectively identify the unknown system and observation noise variances. In this paper, we present the same normalization technique for nonlinear KF and RTSS algorithms named the unscented KF and unscented RTSS. In the same way as the normalized KF and normalized RTSS, these new normalized unscented KF and normalized unscented RTSS algorithms make it possible to estimate the unknown noise variances of nonlinear state-space models. Because it often happens that the noise variances are unknown in actual analysis cases, these algorithms are considerably effective from the aspect of the application viewpoints. The performance was confirmed by experiments using artificially generated data.

I. INTRODUCTION

This paper presents a normalized unscented Kalman filter (NUKF) and a normalized unscented Rauch-Tung-Striebel smoother (NURTSS) for nonlinear state-space model identification. Before describing the algorithms, we first briefly review linear and nonlinear state-space models corrupted by Gaussian noises and the widely-used filter and smoother algorithms. A linear state-space model is described by the following equations:

\[
\begin{align*}
X_t &= FX_{t-1} + \mu_{t-1}, \quad \mu_{t-1} \sim N(0, Q) \\
y_t &= HX_t + \omega_t, \quad \omega_t \sim N(0, r)
\end{align*}
\]

(1)

Here, \(X_t\) is a \(L\)-dimensional random vector and \(y_t\) is a scalar observation at time \(t\). \(X_t\) is called a state vector and sequentially estimating it given the observations \(Y_t = \{y_1, y_1, \cdots, y_t\}\) or all of the observations \(Y_N = \{y_N, y_N, \cdots, y_t, y_{t-1}, \cdots, y_1\}\) is the main problem in the filtering and smoothing theories. \(F, H\) are system transition and observation matrices and \(\mu_{t-1}, \omega_t\) are system and observation noises following Gaussian distributions both with zero means and a variance-covariance matrix \(Q\) and a variance \(r\), respectively. Suppose that the model design parameters \(F, H, Q, r\) are all known, the KF[1] and RTSS[2] provide the minimum mean square estimates given \(Y_t\) or \(Y_N\).

A nonlinear state-space model corrupted by Gaussian noises is described in the following equations:

\[
\begin{align*}
X_t &= f(X_{t-1}, \mu_{t-1}), \quad \mu_{t-1} \sim N(0, Q) \\
y_t &= HX_t + \omega_t, \quad \omega_t \sim N(0, r)
\end{align*}
\]

(3)

(4)

Here, \(f(\cdot)\) is a nonlinear function and because of this nonlinear transformation, even though the system and observation noises both follow Gaussian distributions, the state vector \(X_t\) no longer follows a Gaussian distribution. Unfortunately, generally speaking, filters or smoothers that provide the minimum mean square estimates do not exist for nonlinear models and thus, several approximation filters and smoothers have been invented. The UKF[3] and URTSS[4] are the widely used algorithms because of their high estimation accuracy also with the computational efficiency. UKF and URTSS algorithms are described in the following equations:

\[
\begin{align*}
\chi_0^* &= X_{t-1|t-1}^* \\
\chi_j^* &= X_{t-1|t-1}^* + \left(\sqrt{L^* + \lambda}V_{t-1|t-1}^* \right) j \\
L^* + \lambda &= \alpha^2(L^* + \kappa) \\
w_j^e &= \left\{ \frac{\lambda}{(L^* + \kappa)}, \right. \quad \text{for } j = 0 \\
&\left. \frac{\beta}{(2(L^* + \kappa))}, \right. \quad \text{for } j \neq 0 \\
w_j^c &= \left\{ \frac{\lambda}{(L^* + \kappa)} + (1 - \alpha^2 + \beta), \right. \quad \text{for } j = 0 \\
&\left. \frac{\beta}{(2(L^* + \kappa))}, \right. \quad \text{for } j \neq 0 \\
X_{t|t-1} &= \sum_{j=0}^{2L^*} w_j^e f(\chi_j^*) \\
C_{t|t-1} &= \sum_{j=0}^{2L^*} w_j^c \left( X_j - X_{t-1|t-1} \right) \\
&\times f(\chi_j^*) - X_{t|t-1}^T \\
V_{t|t-1} &= \sum_{j=0}^{2L^*} w_j^c \left( f(\chi_j^*) - X_{t|t-1} \right) \\
&\times (f(\chi_j^*) - X_{t|t-1})^T \\
K_t &= V_{t|t-1} H^T (HV_{t|t-1} H^T + r)^{-1} \\
X_{t|t} &= X_{t|t-1} + K_t (y_t - HX_{t|t-1}) \\
V_{t|t} &= V_{t|t-1} - K_t H V_{t|t-1}
\end{align*}
\]

(5)

(6)

(7)

(8)

(9)

(10)

(11)

(12)

(13)

(14)

(15)

†Masaya Murata and Kunio Kashino are research scientists at NTT Communication Science Laboratories, NTT Corporation, 3-1, Morinosato Wakamiya, Atsugi-Shi, Kanagawa 243-0198, Japan. murata.masaya at lab.ntt.co.jp

‡Kunio Kashino is also a visiting professor at National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430 Japan.
\[
\begin{align*}
\text{URTSS} \\
D_t &= C_{t+1|t} V_{t+1|t}^{-1} \quad (16) \\
X_{t|N} &= X_{t|t} + D_t (X_{t+1|N} - X_{t+1|t}) \quad (17) \\
V_{t|N} &= V_{t|t} + D_t (V_{t+1|N} - V_{t+1|t}) D_t^T \quad (18)
\end{align*}
\]

Here, \(X_{t-1|t-1} \) and \(V_{t-1|t-1} \) are the augmented state vector and the augmented variance-covariance matrix of \(X_{t-1|t-1} \) and \(V_{t-1|t-1} \) as defined below:

\[
\begin{align*}
X_{t-1|t-1} &= \begin{bmatrix} X_{t-1|t-1}^T \end{bmatrix} \quad (19) \\
V_{t-1|t-1} &= \begin{bmatrix} V_{t-1|t-1} & 0 \\ 0 & Q \end{bmatrix} \quad (20)
\end{align*}
\]

\(X_{t-1|t-1}, V_{t-1|t-1} \) are the expectations and the variance-covariance matrices of the filtered state at time \(t - 1\), given \(Y_{t-1} = \{y_{t-1}, y_{t-2}, \ldots, y_1\} \). \(X_{t|t-1}, V_{t|t-1} \) are the expectations and the variance-covariance matrices of the predicted state at time \(t\), given \(Y_{t-1}, K_t \) and \(D_t \) are Kalman gain and smoother gain at time \(t\). \(L^*\) is the dimension of the augmented state vector \(X_{t-1|t-1} \). \(\chi, \alpha, \beta\) are UKF design parameters and they are usually set to 3, 10^{-3}, 2, respectively. \(\sqrt{(L^* + \lambda)V_{t-1|t-1}}\) denotes the \(j\)th column of Cholesky square root matrix of \((L^* + \lambda)V_{t-1|t-1}^{-1}\), and \{\(\chi, w^*, u^*\}\} are 2\(L^* + 1\) number sigma points generated by using the columns and their corresponding weights. \{\(\chi\}\) denote the sigma points discarding the elements of the system noise. Because the observation eq. (4) is linear, the Kalman gain and the filtering procedures are the same as those of the KF. \(C_{t+1|t}\) is the predicted cross covariance between the state estimates at time \(t\) and \(t + 1\). The unscented transformation approximation of the nonlinearly transformed state based on the sigma points is accurate up to at least second order in terms of the Taylor series expansion around the posterior estimate of the state. Therefore, even though the UKF and URTSS are no longer optimal, for strong nonlinearity cases, they provide more accurate state estimates than those obtained by using the extended KF\(^5\) or linearized RTSS. In addition, the algorithms do not require the burden-some Jacobian calculations and are capable of handling the discontinuities in the nonlinear function.

These algorithms erected monumental landmarks from the viewpoints of estimation accuracy and computational efficiency in optimal state estimation theory. Furthermore, because they are easy to understand and implement, the KF, RTSS, UKF, and URTSS are widely used in many application fields from physical and engineering system identifications to economic time series analysis. However, the problem is that these algorithms only work under the condition that the system and the observation models are well identified. This assumption is rather too strong and indeed in actual analysis cases, it often happens that the noise statistics are unknown. For example, when the system model is not correctly specified, the corresponding system noise, which is equivalent to the modeling error brought to this model, becomes large and unknown. For such dynamic systems, the state-of-the-art filtering and smoothing algorithms described above no longer guarantee sufficient accuracy for the state estimation results.

However, for the linear state-space model, initially normalizing KF and RTSS algorithms by the total of the unknown system and observation noise variances yields new state estimation algorithms and they are applicable to the unknown model identification problem. These algorithms are called the normalized KF and normalized RTSS. Our proposed algorithms are the extensions to nonlinear cases and we call them the normalized UKF and normalized URTSS. By applying these algorithms to nonlinear state-space models, the unknown noise variances are estimated with high precision and the models are effectively identified.

The paper is organized as follows. We have introduced the concepts of linear and nonlinear state-space models and reviewed the widely used state estimation algorithms in section I. In section II we describe the normalized KF, RTSS, UKF, and URTSS to cope with the identification problems for linear and nonlinear state-space models with unknown Gaussian noises. In section III we show the experimental results using various artificially generated data and show the effectiveness of these algorithms. Section IV describes the related research and in section V, we conclude the paper.

II. NORMALIZATION BY TOTAL NOISE VARIANCES

Normalized KF and normalized RTSS are the state estimation algorithms for a linear state-space model corrupted by additive Gaussian noises. We first explain the algorithms and the model identification method using the state estimation results. We next describe the normalized UKF and normalized URTSS algorithms for nonlinear state-space models with unknown Gaussian noises. We then present the nonlinear model identification method based on the iteration of the algorithms. The dimensions of the following state vectors are all assumed to be 1.

A. Normalized KF and Normalized RTSS

Recall the KF and RTSS algorithms. Dividing the variance-covariance matrices of the state vectors by a certain value yields the same expectation estimates of the state vectors. This is easily confirmed by calculating the first filtering step as shown below:

\[
\begin{align*}
\tilde{V}_{0|0} &= V_{0|0} / c, \quad Q' = Q / c, \quad r' = r / c \\
\tilde{V}_{1|0} &= F V_{1|0} F^T + Q' \\
K_1 &= V_{1|0} H^T (H V_{1|0} H^T + r')^{-1} \\
&= V_{1|0} H^T (H V_{1|0} H^T + r)^{-1}
\end{align*}
\]

Here, \(c\) is a normalization number. The above equations show that the original KF and the normalized KF estimate the same Kalman gain value \(K_1\) for the first filtering step. Therefore, the expectations of the filtered states obtained by using these algorithms perfectly agree. In the same way, the subsequent Kalman gains \(K_2, K_3, \cdots, K_t\) and the state expectation estimates \(\tilde{X}_{2|2}, \tilde{X}_{3|3}, \cdots, \tilde{X}_{t|t}\) all become the
same. Like the filtering, the smoothing gain $D_t$ also becomes
the same as the original RTSS as shown below:

$$D_t = V_{t|t}^r F^T V_{t+1|t}^{r-1} = V_{t|t} F^T V_{t+1|t}^{-1}$$

Therefore, the normalization of the algorithm yields the same
expectation estimates of the smoothed state vectors.

Algorithms of the normalized KF (NKF)[6] and the
normalized RTSS (NRTSS) utilize the estimation consistency of the
state expectations and replace the normalization number
with the total of the unknown noise variances $\nu = q + r$. The
normalizations of $V_{t-1|t-1}, q, r$ by $z$ are as follows:

$$V_{t-1|t-1}' = V_{t-1|t-1}/z$$
$$q' = q/z = \gamma$$
$$r' = r/z = 1 - \gamma$$

Here, $\gamma$ is a newly introduced parameter denoting the ratio
of the system noise variance to the unknown total variance.
Replacing $V_{t-1|t-1}, q, r$ with $V_{t-1|t-1}', q' = \gamma, r' = 1 - \gamma$
yields the algorithms of the NKF and NRTSS summarized
in the following equations:

**NKF**

$$X_{t|t-1} = FX_{t-1|t-1}$$
$$V_{t|t-1}' = FV_{t-1|t-1}'F^T + \gamma$$
$$K_{t} = V_{t|t-1}'(HV_{t|t-1}' - 1)^{-1}$$
$$X_{t|t} = X_{t|t-1} + K_{t}(y_{t} - HX_{t|t-1})$$
$$V_{t|t}' = V_{t|t-1}' - K_{t}HV_{t|t-1}'$$

**NRTSS**

$$D_t = V_{t|t}' F^T V_{t+1|t}^{r-1}$$
$$X_{t|t}' = X_{t|t} + D_t(X_{t+1|N} - X_{t+1|t})$$
$$V_{t|t}' = V_{t|t}' + D_t(V_{t+1|N} - V_{t+1|t})D_t^T$$

The linear model identification method based on the state
estimation results uses the estimate of the total variance $\hat{\nu}$
calculated by the following equations:

$$\nu_N \equiv y_N - HX_{N|N-1}$$
$$\hat{\nu} = V[\nu_N]/(HV_{N|N-1} - 1)^T + (1 - \gamma)$$

Varying the parameter $\gamma$ from 0 to 1 with a small interval
e.g. 0.1, $\gamma$ scoring the largest $I(\gamma)$ value is adopted as the
ratio estimate denoted by $\hat{\gamma}$. Then, the variance estimates of
the system noise $q$ and the observation noise $r$ are calculated by
$\hat{q} = \hat{\nu} \hat{\gamma}$ and $\hat{r} = (1 - \hat{\gamma}) \hat{\nu}$, respectively. The essential
point is that without knowing the true total variance of the
unknown noises, the NKF and NRTSS are executable with the
setting of $\gamma$. Then the linear model is identified with the
estimates of unknown noise variances.

**B. Normalized UKF and Normalized URTSS**

The similar normalized algorithms are desired for the identification
of nonlinear state-space models corrupted by unknown Gaussian Noises.
However unlike the linear case, nonlinear transformations prohibit us from applying the same
normalization procedure as described in the previous section.
Because of the nonlinear effects, aforementioned normalizations yield a different expectation estimate of the predicted
state $X_{t|t-1}$ from that obtained by the UKF. However, this inconsistency of the resulting expectation estimates is rescued by using the following sigma points estimation of the predicted state expectation.

$$\chi^{*}_{0} = X_{t-1|t-1}$$
$$\chi^{*}_{j} = X_{t-1|t-1} + \pm \sqrt{V^{*}_{t-1|t-1}}$$
$$w^{*}_{j} = \left\{ \begin{array}{ll}
\frac{-L_{j} + 1}{2}n_{j}^{2}, & \text{for } j = 0 \\
\frac{1}{2n_{j}^{2}}, & \text{for } j \neq 0
\end{array} \right.$$

Here, $X_{t-1|t-1}^{*}$ is the augmented state vector defined by
(19) and $V_{t-1|t-1}^{*}$ is the normalized augmented variance-
covariance matrix defined as below,

$$V_{t-1|t-1}^{*} = \begin{pmatrix}
V_{t-1|t-1}/z & 0 \\
0 & q/z
\end{pmatrix}$$

The above estimation algorithm is obtained by choosing
$\alpha^2 (L^* + \kappa) = 1/z$ in the original UKF algorithm instead of setting them $1 \times 3 = 3$. Because $\alpha$ is arbitrary and is
just recommended to be a small value for the case of $f(\cdot)$
showing strong nonlinearity, the sigma points and the weights
parametrized by $\gamma$:

$$V_{t|t-1}^{*y} = HV_{t|t-1}^{*y} T + (1 - \gamma)$$
$$I(\gamma) = -\frac{1}{2} \left\{ N \log \pi + \sum_{i=1}^{N} \log (\hat{\nu} V_{t|t-1}^{*y})
+ \sum_{i=1}^{N} (\nu_i)^2 (\hat{\nu} V_{t|t-1}^{*y})^{-1}\right\}$$

$$= -\frac{1}{2} \left\{ N \log \pi + N \log \hat{\nu} + \sum_{i=1}^{N} \log (\hat{\nu} V_{t|t-1}^{*y})
+ \frac{1}{z} \sum_{i=1}^{N} (\nu_i)^2 (\hat{\nu} V_{t|t-1}^{*y})^{-1}\right\}$$

(32)
\( \chi^*, w^s \) generated from the normalized variance-covariance matrix \( V^*_t \) correspond to those generated from \( V^*_{t-1} \) with \( L^* + \lambda = 1/z \) setting. Thus, the sigma point estimations of \( X_{t-1} \) obtained by using the above algorithm and by using the UKF algorithm are exactly the same.

To construct the normalized algorithms, we define the new weights for the variance-covariance matrix calculation to obtain \( V^*_t = V^*_t / z \) as below:

\[
w^s_j = \begin{cases} 
\frac{w^s_j + (1 - \alpha^2 + \beta)}{\alpha}, & \text{for } j = 0 \\
\frac{w^s_j}{\alpha}, & \text{for } j \neq 0 
\end{cases}
\]

\[
V^*_t = \sum_{j=0}^{2L^*} w^s_j \left( f(X^*_{t-1}) - X^*_{t-1} \right) \times \left( f(X^*_{t-1}) - X^*_{t-1} \right)^T
\]

Here, \( \alpha = \sqrt{1/(3z)} \) because \( (L^* + \lambda) \) in (7) is set to be \( 1/z \) for the normalization consistency (recall \( L^* + \kappa = 3 \)). These time update algorithms produce the same Kalman gain \( K_t \) as that gained by the UKF. Thus, the subsequent filtering procedure yields a state expectation estimate identical to that obtained by the UKF. Consequently, the subsequent Kalman gains \( K_2, K_3, \ldots, K_t \) and the state expectation estimates \( X^*_{2t}, X^*_{3t}, \ldots, X^*_{tt} \) also become the same. The normalized cross-covariance matrix \( C^*_t \) is estimated in the same way.

As well as the normalizedKF and RTS, replacing \( V^*_{t-1} \), \( q \), \( r \) with \( V^*_t \), \( q' = \gamma \), \( r' = 1 - \gamma \) yields the algorithms of normalized UKF (NUKF) and normalized URTSS (NURTSS) summarized in the following equations:

**NUKF**

\[
\begin{align*}
\chi^*_0 &= X^*_0 \\
\chi^*_j &= X^*_{t-1} + \left( \sqrt{V^*_t} \right)_j \\
w^s_j &= \begin{cases} 
\frac{-L^* + 1/z}{1/z}, & \text{for } j = 0 \\
\frac{-1}{\sqrt{V^*_t}}, & \text{for } j \neq 0 
\end{cases} \\
w^c_j &= \begin{cases} 
\frac{w^s_j + (1 - \alpha^2 + \beta)}{\alpha}, & \text{for } j = 0 \\
\frac{w^s_j}{\alpha}, & \text{for } j \neq 0 
\end{cases} \\
X^*_{t-1} &= V^*_t + X^*_{t-1} - X^*_{t-1} \\
C^*_t &= \sum_{j=0}^{2L^*} w^c_j (X^*_j - X^*_{t-1}) \times \left( f(X^*_{t-1}) - X^*_{t-1} \right)^T \\
V^*_t &= \sum_{j=0}^{2L^*} w^c_j (f(X^*_j) - X^*_{t-1}) \times \left( f(X^*_j) - X^*_{t-1} \right)^T \\
K_t &= V^*_t H^T (HV^*_t - (1 - \gamma))^{-1} \\
X^*_{t+1} &= X^*_{t+1} - K_t (y_t - HX^*_{t+1}) \\
V^*_t &= V^*_t - K_t H V^*_t
\end{align*}
\]

**NURTSS**

\[
\begin{align*}
D_t &= C_t V^*_t^{-1} \\
X^*_{t+1} &= X^*_{t+1} + D_t (X^*_{t+1} - X^*_{t+1}) \\
V^*_t &= V^*_t - X^*_{t+1} D_t^T
\end{align*}
\]

The nonlinear model identification method is performed by iterating these algorithms until \( z \) converges to a certain threshold. We denote \( z \) at the \( k \)th time around in the iteration as \( z_k \). Starting from \( z_0 = 1 \), the iteration is repeated until the following convergence criterion is satisfied.

\[
\frac{z_{k+1} - z_k}{z_k} < \eta
\]

Here, \( \hat{z} \) is estimated by (31) and \( \eta \) is the threshold such as 0.001. The criterion is deduced from the theoretical implication that an adequate setting of \( z_k \) approximately holds. After the convergence is satisfied, the resulting \( z_k \) is adopted as the estimate for the total variance of the unknown noises. \( \gamma \) is estimated in the same way as the identification method for linear state-space models. The log-likelihood parametrized by \( \gamma \) is defined as below:

\[
I(\gamma) = -\frac{1}{2} \left\{ N \log 2\pi + N \log z_k + \sum_{t=1}^{N} \log V^*_t \right\} + \frac{1}{2} z_k \left( \mu - \mu_t \right)^2 V^*_t^{-1}
\]

Varying the parameter \( \gamma \) from 0 to 1, \( \gamma \) scoring the largest \( I(\gamma) \) value is identified as the ratio of the system noise variance to the total variance. The estimate \( \hat{\gamma} \) and the resulting variance estimates of the system noise \( \hat{q} = \hat{\gamma} z_k \) and the observation noise variance \( r = (1 - \hat{\gamma}) z_k \) identify the nonlinear state-space model.

**III. EXPERIMENTAL RESULTS**

In this section, we confirm the identification performance of the NKF and NUKF using artificial data. The data are generated from two nonlinear state-space models. The smoothed state estimation results obtained by NURTSS are also presented for comparison with the true state values.

**A. Nonlinear State-Space Model Identification I**

The first experiment addresses the identification problem of a nonlinear state-space model with unknown Gaussian noises. The data is generated according to the following one-dimensional nonlinear system and linear observation models.

\[
x_t = 0.3 x_{t-1}^2 + \mu_{t-1}, \quad \mu_{t-1} \sim N(0, 0.25) \\
y_t = x_t + \omega_t, \quad \omega_t \sim N(0, 0.25)
\]

By changing the samples of the random numbers, we generated 100 different realizations, each including 1000 observation data, with the same initial state \( x_0 = 1 \). The initial settings of the NUKF are \( X^*_0 = 0, V^*_0 = 100, z_0 = 1, \eta = 0.001 \). The identification result by the normalized
extended Kalman filter (NEKF) is also presented for the performance comparison with the NUKF. The NEKF linearizes the nonlinear system model around the posterior estimate of the state $X_{t|t}$. Then the NKF is applied to the linearized model for the identification of the nonlinear state-space model. The extended RTSS (ERTSS) is an RTSS applied to the linearized model and we call the normalized form a normalized ERTSS (NERTSS). The NEKF and NERTSS algorithms are summarized below:

**NEKF**

\[
\begin{align*}
X_{t|t-1} &= f(X_{t-1|t-1}) \\
V'_{t|t-1} &= \tilde{F}V'_{t-1|t-1}\tilde{F}^T + \gamma \\
K_t &= V'_{t|t-1}H^T(HV'_{t|t-1}H^T + (1 - \gamma))^{-1} \\
X_{t|t} &= X_{t|t-1} + K_t(y_t - HX_{t|t-1}) \\
V'_{t|t} &= V'_{t|t-1} - K_tHV'_{t|t-1} \\
\end{align*}
\]

**NERTSS**

\[
\begin{align*}
D_{t} &= V'_{t|t}\tilde{F}^TV'_{t-1|t+1} \\
X_{t|N} &= X_{t|t} + D_{t}(X_{t+1|N} - X_{t+1|t}) \\
V'_{t|N} &= V'_{t|t} + D_{t}(V'_{t+1|N} - V'_{t+1|t})D_{t}^T \\
\end{align*}
\]

Here, $\tilde{F}$ is the Jacobian at $X_{t|t}$ defined as \( \frac{\partial f}{\partial X} \). The above NEKF and NERTSS algorithms yield the same state expectation estimates as those estimated by the EKF and ERTSS. Thus the normalization by the total variance $\gamma$ does not change the estimation results and the estimated $\hat{z}$ is also calculated by (31). The initial settings of the NEKF are $X_{0|0} = 0, V'_{0|0} = 100$.

Here, although the state $x_t$ is one-dimensional, we use the notations $X_{t|t}$ and $V'_{t|t}$ for the correspondence to the algorithms expressed in this paper. For processing each data set, we varied the parameter $\gamma$ from 0 to 1 with a interval 0.1 and $\hat{z}$ scoring the largest $I(\gamma)$ value is adopted as the estimated ratio $\hat{z}$ of the system noise variance to the total noise variances. Then using $\hat{z}, \hat{q}$ and $\hat{r}$ are calculated as the identified system noise and observation noise variances.

The average estimated $\hat{\gamma}$ and the estimated variances $\hat{z}, \hat{q}$, and $\hat{r}$ over the 100 realization data by the NEKF and NUKF are shown in Table I. The identification results obtained by the NEKF contain more errors compared to those obtained by the NUKF. This is because the NEKF approximates the nonlinear function only up to the first order in terms of the Taylor series expansion. On the other hand, because the NUKF is accurate up to the second order in terms of the Taylor series, the overall identification results are superior over those specified by the NEKF.

We depict examples of the identification and the state estimation results for one realization data set in Fig. 1. The NUKF estimated $\hat{\gamma} = 0.5$ and $\hat{z} = 0.47$ and the true states are roughly estimated from the highly contaminated data.

![Fig. 1](image1.png)

**B. Nonlinear State-Space Model Identification II**

The second experiment was performed to test the identification performance under the stronger nonlinearity case. The data is generated according to the following one-dimensional nonlinear system and linear observation models.

\[
\begin{align*}
x_t &= 3\cos(x_{t-1}) + \mu_{t-1}, \quad \mu_{t-1} \sim N(0, 0.25) \\
y_t &= x_t + \omega_t, \quad \omega_t \sim N(0, 0.25)
\end{align*}
\]

We generated 100 different realizations, each including 1000 observation data, with the same initial state $x_0 = 1$. The identification result obtained by the NEKF is also presented for performance comparison. The initial settings of the NEKF and NUKF are the same as those in the previous experiment. The average estimated ratio and the noise variances over the 100 realization data by the NEKF and NUKF are shown in Table II. These results show that for strongly nonlinear state-space models, the NUKF has a potential to yield much superior identification results over the NEKF and, as explained in the previous experiment, it is because that the NUKF approximates the nonlinearly transformed state statistics more accurately than the NEKF does.

We depict examples of the identification and the state estimation results for one realization data set in Fig. 2. The NUKF estimated $\hat{\gamma} = 0.5, \hat{z} = 0.45$ and the true states are estimated with high precision by applying the NURTSS.

![Fig. 2](image2.png)

**TABLE I**

<table>
<thead>
<tr>
<th>(\hat{\gamma})</th>
<th>$\hat{z}$</th>
<th>$\hat{q}$</th>
<th>$\hat{r}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True values</td>
<td>0.5</td>
<td>0.5</td>
<td>0.25</td>
</tr>
<tr>
<td>NEKF</td>
<td>0.55</td>
<td>0.49</td>
<td>0.28</td>
</tr>
<tr>
<td>NUKF</td>
<td>0.47</td>
<td>0.49</td>
<td>0.23</td>
</tr>
</tbody>
</table>

**TABLE II**

| \(\hat{\gamma}\), $\hat{z}$, $\hat{q}$, AND $\hat{r}$ OVER 100 REALIZATION DATA. |
|--------|------|------|------|
| True values | 0.5 | 0.5 | 0.25 | 0.25 |
| NEKF | 0.75 | 0.80 | 0.62 | 0.18 |
| NUKF | 0.51 | 0.54 | 0.31 | 0.22 |
The overall identification and state estimation results presented in this section confirm the effectiveness of the NUKF and NURTSS. In this paper, we assumed that the state-space models were composed of nonlinear system models and linear observation models. However, the extension to nonlinear observation models are straightforward. Generating sigma points from $V^{T}_{U(t-1)}$ and estimating the statistics of the prediction using the weights in (36) and (37) yields the same Kalman gain as that obtained by the UKF. Thus, the resulting state expectation estimates also become the same as those obtained by the UKF and the identification method presented in this paper can be applied to estimate the unknown parameters of the nonlinear state-space model.

IV. RELATED RESEARCH

The problem of Kalman filter under unsure state-space models has been investigated by several authors[7][8][9][10][11]. For a linear state-space model with unknown additive Gaussian noises, the model identification method was pioneered by Mehr[a][12][13]. For a nonlinear state-space model, the multiple-model estimation method provided by Simon[14] may be effective in the identification problem. The method uses $N$ different model design parameter sets (that is, it uses $N$ different state-space models), and runs the $N$ parallel KFs simultaneously. Because the concept of the method is not limited to linear state-space models, the multiple-model estimation technique is promising also for identification problems of nonlinear state-space models.

The filtering techniques under unsure state-space models are called adaptive filtering or robust filtering[15]. Among them, the fading-memory filter[16] and $H_\infty$ filter[17] are widely used. The fading-memory filter gives greater emphasis to more recent data and forgets the past state estimations more rapidly than a KF. Because more uncertainty is imposed on a priori predicted state, the state to be filtered becomes more sensitive to the received observation. In other words, the filter is designed to be equipped with robustness to the unsure system model.

Although the $H_\infty$ filter has a robust feature similar to that of the fading-memory filter, the concept is totally different. The $H_\infty$ filter estimation assumes that the noise is worst-case and minimizes such worst-case estimation error while the KF just minimizes the estimation error under the given magnitudes of noises. Because accurate system models are not as readily available for real-world phenomena, the concept of the $H_\infty$ filter seems to be more suitable. However, many design parameters have to be pre-specified before the execution and its extension to nonlinear state-space models is not yet well-understood. Further research with investigation of the unknown parameter identification methods is highly desired. The mixed Kalman/$H_\infty$ filter[18] that combines the best features of Kalman filtering with the best features of $H_\infty$ filtering may also be promising for the identification problem of unsure state-space models.

V. CONCLUDING REMARKS

In this paper, we presented new state estimation algorithms, called a normalized unscented Kalman filter (NUKF) and normalized unscented Rauch-Tung-Striebel smoother (NURTSS), to identify unknown system noise and observation noise variances of nonlinear state-space models.

In the analyses of the real-world phenomena, the dynamics of the true states are mostly unknown and just applying the state-of-the-art optimal state estimation techniques such as the KF and RTSS to such uncertain models can lead to the large estimation errors. The problem becomes considerably serious when the models are expressed in terms of the nonlinear functions of the states. Our proposed algorithms and methods are capable of estimating the states along with the identification of unknown variances of system and observation noises brought to the nonlinear state-space models.

Our future works include the multi-dimensional state estimation problems. We will also investigate the capabilities of our algorithms for real-world identification problems.

REFERENCES